NORMAL HUMAN AGING:

The Baltimore Longitudinal Study of Aging

NATHAN W. SHOCK

Richard C. Greulich
Reubin Andres
David Arenberg
Paul T. Costa, Jr.
Edward G. Lakatta
Jordan D. Tobin

U. S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Institutes of Health
National Institute on Aging
Gerontology Research Center
Baltimore City Hospitals

NIH Publication No. 84-2450
November, 1984

CHAPTER I

Methods for the Study of Aging

The two principal methods by which the effect of aging on a variable can be measured are the cross-sectional and longitudinal designs. The cross-sectional method is characterized by measurements made at approximately the same time on a large number of subjects covering the entire adult age span. Age changes are not measured directly but are inferred from the differences in mean values observed in different age groups or from the overall regression of the measurement on age. Only average differences between age groups are identified.

The longitudinal method is characterized by serial measurements of a specific variable on the same subject as aging occurs; it thus identifies age changes in individuals in addition to average differences between groups of subjects of different ages. Since each method has its strengths and weaknesses, the quantitative measurement of aging requires the application of both. Other variations of the two ("cross-sequential" and "time-sequential" approaches) are needed to control factors that cannot be isolated by either method alone.

THE CROSS-SECTIONAL METHOD AND ITS LIMITATIONS

1. Advantages and Limitations

Growth and development have traditionally been studied by the cross-sectional method, by which the average values of a variable are calculated for groups of subjects distributed according to age. Growth is inferred from the progressive increase of the average values for height or weight in groups of growing children; that is, the regression of measurements on age is viewed as an index of growth rate. Such analysis neither provides a direct measurement of age changes nor specifies the magnitude or rates of change in individual subjects. Its primary advantage is that the presence of age trends in a group of subjects can be detected fairly quickly. Caution is necessary in its interpretation, however, since differences between age groups include birth-cohort[1] as well as age effects (see below, "Strategies of Analytical Design").

Students of child development recognized this limitation of cross-sectional analyses in the early part of this century and initiated longitudinal studies, in which measurements of height and weight were repeated at short intervals to generate growth curves for individual subjects (Dearborn et al., 1938). Although important findings have emerged, the longitudinal method has been used in only a few studies of adults (see Chapter II), because of the difficulties of recruiting and retaining subjects for repeated measurements over long periods of time, as well as of finding the necessary long-term financial support.

---

[1]A birth cohort consists of individuals born in the same arbitrarily chosen interval of time. Since specific environmental conditions occur at different ages for subjects from different birth cohorts, the effects of such events may be confounded with the effects of aging in their influence on cross-sectional measurements.

2. Effects of Differential Mortality

As a study population ages it becomes more and more selected, since death occurs more frequently among old than among young subjects. By age 70, the population available for study represents only about 50% of the original birth cohort. Averages derived from measurements in young adults are thus based on observations of subjects some of whom will not live to age 70, while data from older subjects obviously represent an "élite" population that has survived.

Age differences in a measured variable do not necessarily reflect changes in individuals or even the average changes in specific age groups, since deaths do not occur randomly throughout the population but are more likely among individuals whose characteristics increase their susceptibility. The selective nature of the process is implied by the term "differential mortality." This effect increases with age.

The influence of differential mortality on inferences about age changes made on the basis of cross-sectional measurements of age differences can be visualized from the theoretical curve in Figure 1. Consider the hypothetical variable X whose level varies among individuals but does not change with aging in any individual. Each of the horizontal lines in the figure represents a single individual as he ages. Suppose further,

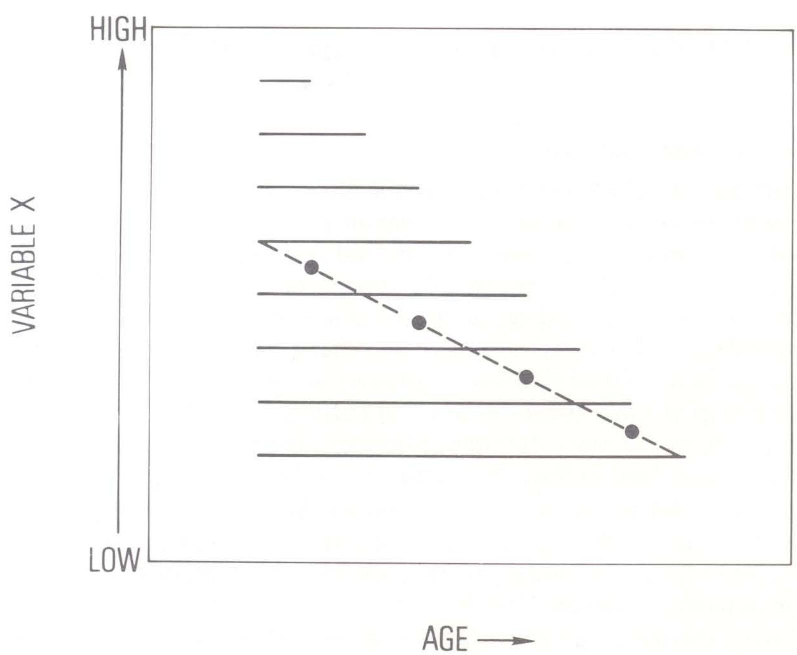

Figure I.1. Confounding effect of selective mortality on inferences about age changes. Each solid line segment represents the pattern of change in an individual (in this case there is no change with aging, and death occurs at the end of the line). High values for the variable X are assumed to be deleterious. The closed circles represent mean values that would be obtained in a cross-sectional study; the dashed line connecting these dots would then correctly represent age *differences* among age groups, but the inference that age *changes* were occurring in individuals would be erroneous.

as the figure shows, that a high value of X is deleterious, so that deaths occur first in high-X and last in low-X subjects. As the figure indicates, the mean value for X at the earliest ages will be high, since all subjects are alive; at each succeeding decade the mean will fall, since the subjects with the highest values will have been eliminated by death. As a result, the circles connected by the dashed line represent the average values that would be obtained from a single cross-sectional study. The average values of X clearly fall with advancing age, although within individual subjects X does not change with age.

Figure 2 illustrates another way in which cross-sectional analysis may lead to erroneous conclusions. One can imagine a function (variable Y) that declines linearly across age groups but in which a floor effect, or "lethal limit," appears. In this case a cross-sectional study might show an average decline that would not accurately represent the magnitude of age changes in individuals. It is even possible (Fig. 3) to picture a variable that declines with age but in which younger subjects can tolerate lower levels than older subjects before dying. In this case the cross-sectional analysis might show no age differences at all despite a decline in function in individual subjects across the age span.

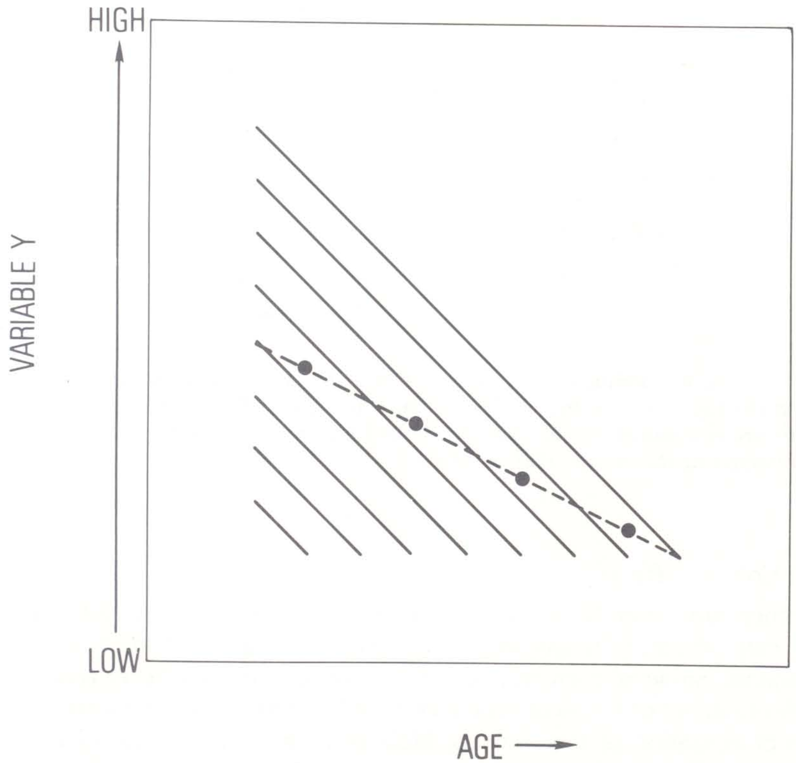

Figure I.2. Confounding effect of selective mortality on the magnitude of age changes. See Fig. 1 for explanation of line segments. A floor effect or lethal limit is assumed for the variable Y. In this case the dashed line representing age *differences* would underestimate true age *changes*.

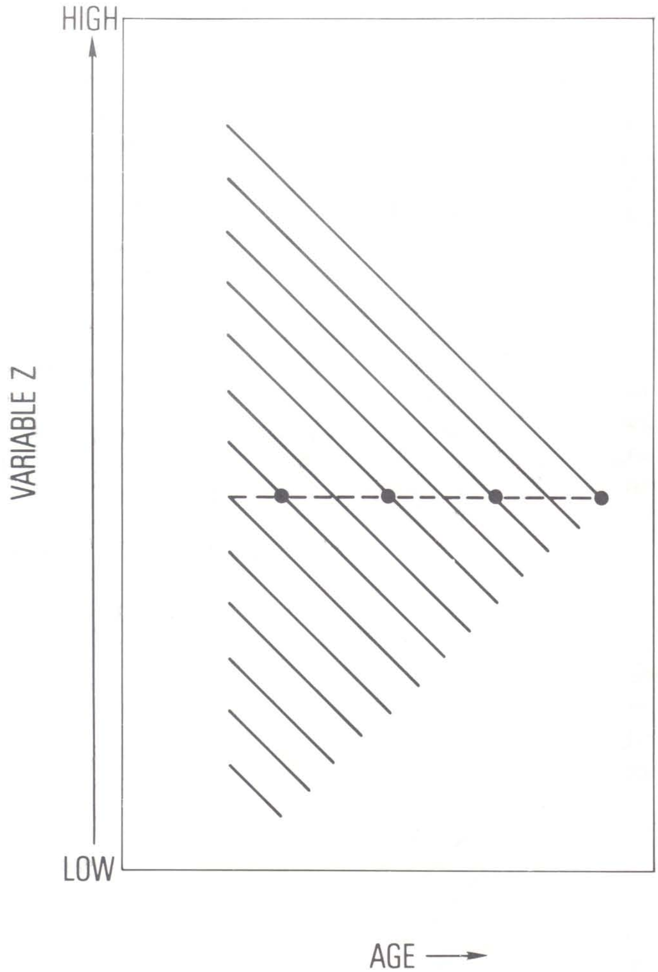

Figure I.3. Confounding effect of selective mortality as a result of which age changes are not revealed in cross-sectional studies. See Fig. 1 for explanation of line segments. A lethal limit that varies with age is assumed. In this case the dashed line would show no age *differences*, although large age *changes* in individuals had occurred.

3. Birth-Cohort Effects

Another limitation of cross-sectional studies is that young and old groups of subjects may differ in characteristics other than age that may also affect the measurements. Subjects born in a specified span of calendar years represent a birth cohort. An example of birth-cohort effects may be found in tests that are influenced by the level of education of the subjects. Most of today's young adults have completed high school, while a much smaller proportion of adults educated in the early years of this century reached that level of education. In any test in which level of education influences performance, young subjects will thus out-perform the old—but the difference may be due to education rather than to age.

Socioeconomic conditions early in life may also have affected older subjects differently from younger ones. For example, subjects who were 70 years old in 1980 were exposed to the effects of the economic depression of the 1930s when they were in their 20s, whereas subjects born after 1940 have not been subjected to such an event. Similarly, epidemics, wars, and other disruptions that occur at different points of the life cycle of different birth cohorts may influence test results in ways that the cross-sectional method cannot differentiate from true age changes (Birren and Renner, 1977).

4. Disease Effects

Since the occurrence of many diseases increases with advancing age, one of the primary problems in attributing differences between groups of subjects to aging is the necessity of excluding subjects suffering from diseases that influence the variable under study. This is an extremely difficult problem for which there is no certain remedy.

Many investigators, especially those concerned with behavioral and social research, have simply ignored the problem. Others have set criteria of health status, ranging from superficial to comprehensive, for inclusion in the study. These are commonly limited to the identification of a few specific disease states such as coronary artery disease or diabetes, or to arbitrary standards of normality in physical findings or laboratory tests such as blood pressure, blood glucose, or hemoglobin concentration. In many studies health status was determined by self-evaluation of the subjects; if a subject said he was in good health he was regarded as healthy. In only a few instances was a detailed physical examination carried out by physicians to screen subjects for the presence of specific diseases.

THE LONGITUDINAL METHOD

Some of the limitations of cross-sectional studies of aging can be minimized or overcome by a longitudinal design, in which the same subjects are measured repeatedly. The ideal longitudinal study of aging would provide observations on individuals over their entire life spans. Since this design is impractical, most longitudinal studies have been limited to specific periods of the life cycle. The phases of growth and development, for example, have received much more attention than adult aging. Among the many critical questions about how adults age that can be answered only from serial observations are the following:

Does the average curve of age differences based on cross-sectional data represent the average progression of aging in individual subjects?

How rapidly does an individual change with respect to a specific variable or test? That is, what is the rate of change? What is the diversity among individuals?

Is there a general aging factor, or does each organ system show a different pattern of aging? How are age changes in different variables related in individual subjects?

Do critical events in the life cycle of an individual affect aging? An answer to this question requires serial measurements in the subject before and after the event, which may take such forms as a heart attack or other severe illness, exposure to toxic substances or radiation, loss of job, retirement, loss of spouse, loss of mobility, or loss of independent living.

Can patterns or levels of performance at a given age be used as predictors of performance at a later age, or of longevity?

Can aging be distinguished from disease?

Does age influence the progression of such disease states as diabetes, arteriosclerosis, and hypertension?

Can a causal ordering be determined from serial observations when two variables are known to co-vary?

ADVANTAGES OF THE LONGITUDINAL METHOD

1. Age Regressions for Individual Subjects

The primary advantage of the longitudinal method is that it makes it possible to estimate age changes in an individual over a specified period, so that a "rate of aging" may be determined for any specified variable. The study must be so designed that enough observations are made in an individual to permit calculation of the standard error of estimate of the calculated regression on age. It is usually assumed that the regression of the variable on age is linear. Although the assumption of linearity may or may not be true, it is seldom possible to collect enough observations to reject it. A statistical analysis *(Schlesselman, 1973a,b)*[2] has been made of the experimental strategies (duration of the study, frequency of observations, number of observations) essential to achieve a specified reliability of estimate of the individual regression slopes for a specific variable. For one variable, equally satisfactory slopes might be computed by carrying out three tests in 14 years, ten tests in ten years, or 30 tests in six years. For another it might not be possible to compute individual slopes with satisfactory accuracy unless monthly examinations were carried out for 30 years. Two characteristics of the variable that necessitate different planning strategies are the mean rate of change with time (the mean regression on age for the population) and the degree of variance in the individual slopes. A third factor is the investigator's determination of the degree of accuracy required in the estimate of the age regression for an individual subject. The investigator may thus choose among many strategies for the design of a longitudinal study.

2. Predicting Outcomes

Events or processes experienced at various times may affect a person's health or functioning in later life, as well as survival. The longitudinal study design is valuable, sometimes essential, to identify events of significance and to quantify their long-term effects. In one sense, a true prospective longitudinal study with repeated periodic evaluations is not required to answer such questions: Information concerning past events may be obtained by history, and outcomes may be sought at a chosen point. The value of such an analysis, however, may be limited by the inaccuracies inherent in historical recall of distant events. The longitudinal approach, with reasonably frequent evaluations, decreases memory error, may provide objective evidence for the presence of an event, and permits more accurate identification of the time when both the event and its effects occur.

[2]References that appear in italics in the text indicate longitudinal studies on the BLSA population. These studies are summarized in Chapter VI and, unless they are still in press, are reprinted in the Appendix.

A major application of longitudinal data is the search for precursors and risk factors related to disease and death. The study of risk factors has traditionally been the province of epidemiological research. The distinctive characteristic of epidemiological research such as the Framingham study (see Chapter II) is its focus on the factors that influence the incidence of disease in a specified human population. One of the strengths of the BLSA lies in the wealth of its measurements, which permits intensive analysis of the antecedents of disease to a degree few epidemiological studies can match.

As a study of aging, the BLSA is also concerned with age-related stability or fluctuation in functions; the multidisciplinary data collected over a period of years can be used to determine predictors of change. Questions that may be addressed include: whether personality differences retard or accelerate declines in cognitive performance; whether regular exercise leads to better pulmonary functioning in old age; and what activities or attitudes contribute to an older person's sense of having lived a satisfying life.

A corollary of the search for significant predictive variables is the identification of their critical levels, the cut-points of which may vary with age. Although 24-hour creatinine clearance, for example, may show a large decline with age, its clinical significance depends on whether the decline predicts an increased likelihood of death or disease. The identification of a critical level or pattern of change may thus be a unique contribution of the longitudinal method.

In addition to assessing the impact of single variables on single outcomes, it is important to test combined effects of variables and complex outcomes. Thus, while myocardial infarction or death must be considered an outcome, more complex outcomes such as the ability to live independently of institutions, the ability to continue the activities of daily living, or the achievement of overall "successful aging" should be assessed. Even the rate of aging of a particular organ system may be analyzed as an outcome (the dependent variable) of other characteristics or risk factors earlier in life.

Studies on children illustrate the power of the longitudinal method in identifying the effects of events that occur at different chronological ages in different subjects. Figure 4A shows a series of curves of individual growth rates derived from serial measurements of height made on children as they aged from five to 18 years (Tanner, 1955). Each individual depicted achieved a maximum growth spurt at a different age, ranging from nine to 14 years. The mean of these curves obtained from cross-sectional measurements (shown by the dotted curve) grossly underestimates the magnitude of the individual adolescent growth spurts and fails to indicate the diversity of their timing. In Figure 4B the same curves have been so arranged that their points of maximum velocity coincide, and other points are plotted as deviations in time from that event. This method can be used to identify the effects of any event on other measurements, provided the time of occurrence of the event can be identified and a series of measurements taken before and after the event is available.

Longitudinal observations also provide an opportunity to identify the effects of physiological events on other variables—effects that cannot be identified by cross-sectional analysis (Shock, 1943). For example, cross-sectional observations of basal heat production in girls and boys aged 11.5 to 17.5 years lead to the conclusion that basal metabolic rate (BMR) falls gradually from age 12 to age 17.5 (Fig. 5) (Shock, 1942). However, when the data for girls are replotted with age of menarche as their zero point (Fig. 6), and serial observations taken at six-month intervals are plotted as

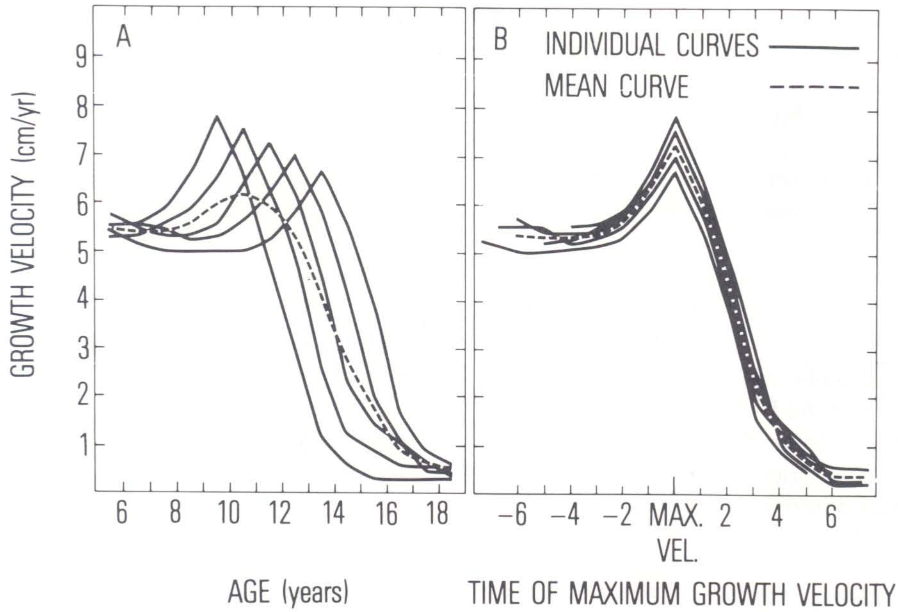

Figure I.4 Relation between individual and mean velocities during the adolescent growth spurt. A: The height curves are plotted against chronological age. B: The height curves are plotted as deviations from time of maximum growth velocity.
From Tanner (1955), after Shuttleworth (1939).

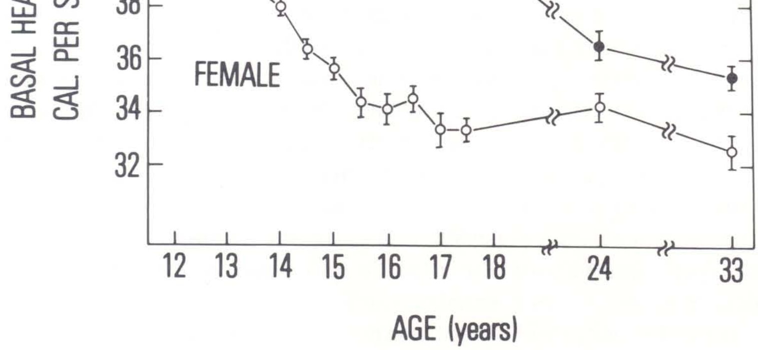

Figure I.5 Average basal heat production in males and females. Measurements made every 6 months from age 11.5 to 17.5 yr in same individuals. Points at ages 24 and 33 obtained from other subjects.
From Shock (1942, adapted).

deviations from age of menarche in each subject, it becomes apparent that menarche—a physiological event—is more important than chronological age in determining the adolescent fall in BMR.

The availability of serial observations also makes it possible to search for the effects of critical events in the life history of individuals—cessation of smoking, death of a spouse, retirement, loss of mobility, need for institutionalization—by comparing measurements made before the event with those made afterward.

3. Continuity of Study Population

The continued availability of a longitudinal population such as that of the BLSA offers unique opportunities for multidisciplinary investigations of the relations between aging and other variables, without the recurring need to recruit new subjects. A major advantage of the BLSA is that it has provided a study population in which thorough and repeated clinical evaluations have been carried out on all subjects. The resultant clinical records add efficiency to the overall operation of the study and provide essential background information on health status to other investigators whose primary interests may lie in domains other than clinical medicine.

Furthermore, when multiple variables are assessed a multiplicative effect occurs, in that each scientist is able to take advantage of the information generated by his colleagues to improve his own subject characterization and research interpretation. As data accumulate in the longitudinal study, not only the quantity but also the quality of the information available on each subject is thus increased.

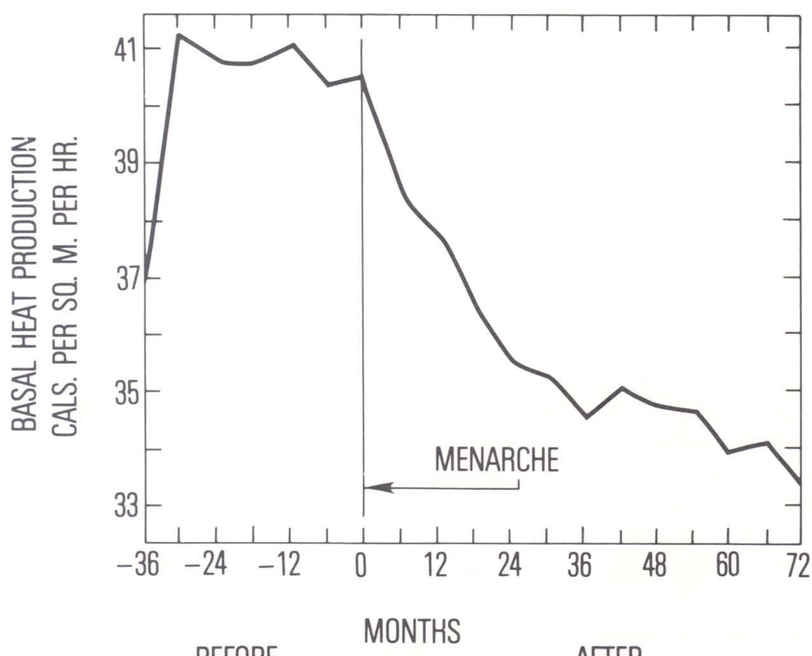

Figure I.6. Effect of menarche on basal heat production. Average values calculated for 6-mo intervals before and after menarche. Zero time for each girl is the age at which she first menstruated.
From Shock (1943).

An inevitable consequence of the existence of a longitudinal panel of subjects in an academic or research environment is that other scientists are attracted by the obvious advantages. This leads not only to a number of spin-off cross-sectional studies but also to the addition of new variables to be studied longitudinally. Any longitudinal study must first go through a phase of cross-sectional analysis. As studies of the initial variables are continued or completed, new ones are introduced; the experimental design evolves into a series of overlapping individual but integrated longitudinal studies.

The addition of new variables five, ten, or 20 years after the beginning of the original study requires subjects across the entire adult age range. Evolution of the study thus mandates the continuing recruitment of new subjects in the youngest age group as time passes and the population ages. The new subjects in turn provide a built-in opportunity for identification of birth-cohort differences and possible period effects within the population (see below, "Strategies of Analytical Design").

OPERATIONAL CHALLENGES IN LONGITUDINAL STUDIES

Although the longitudinal design is essential to the determination of age changes in individuals, it cannot resolve all the difficulties inherent in cross-sectional studies. Furthermore, longitudinal studies have a number of limitations of their own. What seems to be a simple, straightforward question—"How does aging, or the passage of time, affect performance in individual subjects?"—turns out to be a demon in disguise. Many pitfalls in design, subject selection, data collection, and data analysis may undermine or negate the assumption that changes in serially collected measurements are due to aging. These problems include the following:

1. Recruitment and Screening of Subjects

A primary concern in the selection of subjects for a longitudinal study is their commitment to continued participation in the study and their geographic stability over long periods. This requirement limits the sampling procedures that can be used and must be taken into consideration in the generalization of conclusions drawn from any longitudinal study.

The study may require the inclusion of procedures that prove tedious or distasteful to many people. Subjects selected at random will show both a high initial rate of refusal to participate and a high drop-out rate when presented with a test schedule that includes uncomfortable and time-consuming procedures. To this extent most longitudinal studies, including the BLSA, have compromised true representativeness in order to obtain loyalty and cooperation from their subjects.

Some studies from their inception exclude subjects who present clinical or laboratory evidence of disease. Although this procedure may initially limit the study to healthy subjects, it does not avoid the problem of a subject who subsequently develops chronic illness. If the subject is then dropped from the study a great opportunity to trace the historical development of a disease is lost. Hence subjects who developed diseases were not dropped from the BLSA, although observations made on them were no longer included in analyses of age changes.

Another approach is to accept subjects with diagnosed disease, but to exclude measurements made on them from data analyses designed to characterize normal age changes. Serial observations on these subjects as a subgroup can be of great value in distinguishing the effects of aging from those of aging plus disease.

Although the presence of disease will confound the interpretations about aging in both cross-sectional and longitudinal studies, diseases are more apt to be discovered in subjects in a longitudinal than in a cross-sectional study because of the extended time during which longitudinal subjects are seen and tested. Findings that may be equivocal at one testing can be re-examined on subsequent visits for verification of diagnoses.

2. Attrition

Subject losses must be expected as a longitudinal study progresses. Younger subjects are more likely than old ones to move away from the area of the study or to lose interest and motivation. As the subjects become older, death and disability become major factors (Wilson and Webber, 1976). On the other hand, useful research data may emerge from comparison of measurements in subjects who have survived with those in subjects who have died. This may result in development of new methods of predicting the likelihood of death.

Drop-outs due to loss of contact or to subjects' refusal to continue participation pose a problem in the interpretation of results, particularly when it is evident that those who have left the study differed systematically from those who have remained.

The degree to which findings from longitudinal studies are distorted by attrition, whatever its source, depends on the aspect of aging that is being investigated (see Chapter III). Since some variables are influenced more than others by attrition, each variable in each study must be examined for the drop-out effect.

3. Expansion of Subject Panel

Unlike most longitudinal studies, the BLSA is designed to maintain a specified number of subjects within each age decade throughout its course. When new subjects are introduced, it is important that they resemble the original sample as closely as possible. Ideally, this requires careful description and matching of the original and new populations. Although the BLSA did not attempt such a matching, the self-selection strategy employed in the recruitment of its participants has tended to maintain the character of the sample (see Chapter III).

4. Strategies of Analytical Design

It is often assumed that the differences among serial observations within cohorts followed longitudinally represent the effects of aging. This is not necessarily true: A number of non-maturational effects or factors may also induce differences in serial measurements. Changes in measurements made serially over time may be due to: a) changes in procedures; b) systematic methodological error; c) period effects—environmental or cultural changes that may influence all members of the population under study; or d) aging effects.

An aging effect is present if the dependent variable is a function of age regardless of the subject's birth year or of the period or time of observation. A period effect is present if the value of the variable changes systematically as a function of the time of observation and not as a function of age. A birth-cohort effect is present if the value of the variable changes systematically as a function of the subject's birth year rather than of his age.

Table I.1. A Simple Cross-Sequential Design[a]

Date of Birth	Time of Measurement 1960	1970
1900	$\bar{X}_{60}$	$\bar{X}_{70}$
1910	$\bar{X}_{50}$	$\bar{X}_{60}$

[a]Subscripts indicate ages

One of the primary difficulties in the analysis of longitudinal data from a single birth cohort is the confounding of period effects with age changes. Traditional longitudinal designs attempt to circumvent the cross-sectional confounding of aging effects with generational or birth-cohort effects by following the same group of individuals over two or more times of measurement, and thus at two or more ages. Such designs, however, are subject to the confounding of age with period effects. Changes that occur between the first and second measurements may be due to intervening historical events rather than to aging; in some tests, previous exposure or practice may be responsible.

Longitudinal changes include age and period effects. Cross-sectional differences include age and cohort effects. Each set of differences is thus influenced by two of the primary effects, those of age, period, and birth cohort. Cross-sequential and time-sequential designs have been proposed to help untangle the confound. In the cross-sequential design, independent samples of individuals from the same birth cohort are compared at different times of measurement, and thus at different ages (Tab. 1). Since a given individual is measured only once, exposure or practice effects are eliminated. In Table 1 the vertical comparison confounds aging and the effects of birth cohort, while the horizontal comparison confounds aging and period effects.

In the time-sequential design, independent samples of individuals of a specified age are compared at different times of measurement (Tab. 2). Age and time of measurement are separated, but both are confounded with birth cohort. No clear-cut statistical separation of age effects from birth-cohort and period effects can be made.

It was originally assumed that, while each of these designs is ambiguous when used alone, it might be possible to separate out age, period, and birth-cohort effects if all were employed and analyzed simultaneously (Schaie, 1965; Baltes, 1968; Riley et al., 1972; Agnello, 1973; and Mason and Mason, 1973). It has since been demonstrated, however, that there is no single solution to the inevitable confounding of the three, and that interpretation of such analysis depends on the data, the goals of the investigator, and the state of knowledge in the area. *Costa and McCrae (1982)* discuss in greater detail the role of judgment in the interpretation of aging, period, and birth-cohort effects.

5. Maintaining Uniformity of Methods and Quality Control

A longitudinal analysis requires special attention to the maintenance of uniformity of tests and testing conditions throughout the study. Continuous quality control is essential. Methods must be examined at regular intervals for consistency of results and stability of standards.

In the BLSA, replicate samples of blood, urine, and tissues are frozen and stored for re-analysis at a later date. Stored plasma samples have made it possible, for example,

Table I.2. A Simple Time-Sequential Design[a]

Age	Time of Measurement 1960	1970
60	$\bar{X}_{1900}$	$\bar{X}_{1910}$
70	$\bar{X}_{1890}$	$\bar{X}_{1900}$

[a]Subscripts indicate dates of birth

to validate the methodology used for the determination of cholesterol levels by repeating the analyses at one time on a random subset of samples collected over the entire span of the study *(Hershcopf et al., 1982).*

6. Data Storage and Retrieval

Longitudinal studies generate special problems in the storage and retrieval of data. Thanks to the development of computers, it is now possible to store an immense amount of data in such a fashion that the data can be updated as successive test cycles are completed and at the same time remain available for analysis. It is essential that the system and format of data collection be carefully planned in advance, with the advice of personnel trained in computer technology (Ramm and Gianturco, 1974). It is also essential that special precautions be taken to protect the stored data against catastrophic loss as well as to ensure confidentiality.

7. Staffing

Longitudinal studies pose special problems in the recruitment and maintenance of a research staff. As Busse (1965) has pointed out, scientists who participate successfully in longitudinal studies possess distinctive personal characteristics in addition to their scientific qualifications. They must first of all be patient and willing to wait for longitudinal results to evolve. This does not imply that they will sit with folded hands during the early stages of the study; they will have the insight and initiative to examine data cross-sectionally and to look for significant relations among observations as they accumulate. They will generate new hypotheses that can be explored by the introduction of new tests and procedures. As a result, an effective longitudinal study will be dynamic and will not be limited by the initial test procedures.

Since longitudinal studies are apt to be multidisciplinary in their design, the successful investigator should be able to work with others as a member of a team. Each participating scientist should also be interested in other scientific disciplines and willing to communicate with other scientists in the solution of problems.

8. Financing

Longitudinal studies in adults require stable funding for long periods of time if their full potential is to be realized. Although data analysis must be a continuing part of the program, significant longitudinal results cannot be expected in the early years of a study. Hence, a support system that requires the reporting of substantive longitudinal results at short intervals in order to maintain funding is inappropriate. Since short-term funding has in the past characterized most research support, few individuals or institutions have been prepared to initiate and carry out longitudinal studies.

Of primary concern to research administrators are the presumed high costs and the long-term commitment of resources. However, the ability of a longitudinal study to answer certain important questions about aging answerable by no other technique fully justifies the costs. Although costs may appear high in comparison with those of cross-sectional studies, the potential efficiency of having a population with known characteristics available for multiple satellite short-term cross-sectional studies covering the entire period of adult life greatly increases the cost effectiveness of a longitudinal study. The costs of recruiting multiple groups of well-characterized subjects for short-term studies may well exceed those of maintaining a single stable population.